

RECEIVED

SEP 23

TECH CENTER



1600

RAW SEQUENCE LISTING

DATE: 09/16/2003

PATENT APPLICATION: US/09/724,586A

TIME: 16:00:02

Input Set : A:\04871~1.txt

Output Set: N:\CRF4\09162003\I724586A.raw

3 <110> APPLICANT: Sakowicz, Roman
 4 Goldstein, Lawrence
 6 <120> TITLE OF INVENTION: Identification and Expression of Novel Kinesin Motor Protein
 8 <130> FILE REFERENCE: UCSD-04871
 10 <140> CURRENT APPLICATION NUMBER: 09/724,586A
 11 <141> CURRENT FILING DATE: 2000-11-28
 13 <160> NUMBER OF SEQ ID NOS: 9
 15 <170> SOFTWARE: PatentIn version 3.2
 17 <210> SEQ ID NO: 1
 18 <211> LENGTH: 784
 19 <212> TYPE: PRT
 20 <213> ORGANISM: Thermomyces lanuginosus
 23 <220> FEATURE:
 24 <221> NAME/KEY: MISC_FEATURE
 25 <223> OTHER INFORMATION: TL-gamma ATP-dependent plus end-directed microtubule motor
 26 protein
 28 <220> FEATURE:
 29 <221> NAME/KEY: DOMAIN
 30 <222> LOCATION: (1)..(357)
 31 <223> OTHER INFORMATION: kinesin-like microtubule motor domain
 33 <220> FEATURE:
 34 <221> NAME/KEY: DOMAIN
 35 <222> LOCATION: (358)..(442)
 36 <223> OTHER INFORMATION: neck domain links motor domain to stalk domain
 38 <220> FEATURE:
 39 <221> NAME/KEY: DOMAIN
 40 <222> LOCATION: (602)..(784)
 41 <223> OTHER INFORMATION: tail domain
 43 <400> SEQUENCE: 1
 45 Met Ser Gly Gly Gly Asn Ile Lys Val Val Val Arg Val Arg Pro Phe
 46 1 5 10 15
 49 Asn Ala Arg Glu Ile Asp Arg Gly Ala Lys Cys Ile Val Arg Met Glu
 50 20 25 30
 53 Gly Asn Gln Thr Ile Leu Thr Pro Pro Pro Gly Ala Glu Glu Lys Ala
 54 35 40 45
 57 Arg Lys Ser Gly Lys Thr Ile Met Asp Gly Pro Lys Ala Phe Ala Phe
 58 50 55 60
 61 Asp Arg Ser Tyr Trp Ser Phe Asp Lys Asn Ala Pro Asn Tyr Ala Arg
 62 65 70 75 80
 65 Gln Glu Asp Leu Phe Gln Asp Leu Gly Val Pro Leu Leu Asp Asn Ala
 66 85 90 95
 69 Phe Lys Gly Tyr Asn Asn Cys Ile Phe Ala Tyr Gly Gln Thr Gly Ser
 70 100 105 110

ENTERED

RAW SEQUENCE LISTING

DATE: 09/16/2003

PATENT APPLICATION: US/09/724,586A

TIME: 16:00:02

Input Set : A:\04871~1.txt

Output Set: N:\CRF4\09162003\I724586A.raw

```

73 Gly Lys Ser Tyr Ser Met Met Gly Tyr Gly Lys Glu His Gly Val Ile
74      115      120      125
77 Pro Arg Ile Cys Gln Asp Met Phe Arg Arg Ile Asn Glu Leu Gln Lys
78      130      135      140
81 Asp Lys Asn Leu Thr Cys Thr Val Glu Val Ser Tyr Leu Glu Ile Tyr
82 145      150      155      160
85 Asn Glu Arg Val Arg Asp Leu Leu Asn Pro Ser Thr Lys Gly Asn Leu
86      165      170      175
89 Lys Val Arg Glu His Pro Ser Thr Gly Pro Tyr Val Glu Asp Leu Ala
90      180      185      190
93 Lys Leu Val Val Arg Ser Phe Gln Glu Ile Glu Asn Leu Met Asp Glu
94      195      200      205
97 Gly Asn Lys Ala Arg Thr Val Ala Ala Thr Asn Met Asn Glu Thr Ser
98      210      215      220
101 Ser Arg Ser His Ala Val Phe Thr Leu Thr Leu Thr Gln Lys Trp His
102 225      230      235      240
105 Asp Glu Glu Thr Lys Met Asp Thr Glu Lys Val Ala Lys Ile Ser Leu
106      245      250      255
109 Val Asp Leu Ala Gly Ser Glu Arg Ala Thr Ser Thr Gly Ala Thr Gly
110      260      265      270
113 Ala Arg Leu Lys Glu Gly Ala Glu Ile Asn Arg Ser Leu Ser Thr Leu
114      275      280      285
117 Gly Arg Val Ile Ala Ala Leu Ala Asp Met Ser Ser Gly Lys Gln Lys
118      290      295      300
121 Lys Asn Gln Leu Val Pro Tyr Arg Asp Ser Val Leu Thr Trp Leu Leu
122 305      310      315      320
125 Lys Asp Ser Leu Gly Gly Asn Ser Met Thr Ala Met Ile Ala Ala Ile
126      325      330      335
129 Ser Pro Ala Asp Ile Asn Phe Glu Glu Thr Leu Ser Thr Leu Arg Tyr
130      340      345      350
133 Ala Asp Ser Ala Lys Arg Ile Lys Asn His Ala Val Val Asn Glu Asp
134      355      360      365
137 Pro Asn Ala Arg Met Ile Arg Glu Leu Lys Glu Glu Leu Ala Gln Leu
138      370      375      380
141 Arg Ser Lys Leu Gln Ser Ser Gly Gly Gly Gly Gly Ala Gly Gly
142 385      390      395      400
145 Ser Gly Gly Pro Val Glu Glu Ser Tyr Pro Pro Asp Thr Pro Leu Glu
146      405      410      415
149 Lys Gln Ile Val Ser Ile Gln Gln Pro Asp Ala Thr Val Lys Lys Met
150      420      425      430
153 Ser Lys Ala Glu Ile Val Glu Gln Leu Asn Gln Ser Glu Lys Leu Tyr
154      435      440      445
157 Arg Asp Leu Asn Gln Thr Trp Glu Glu Lys Leu Ala Lys Thr Glu Glu
158      450      455      460
161 Ile His Lys Glu Arg Glu Ala Ala Leu Glu Glu Leu Gly Ile Ser Ile
162 465      470      475      480
165 Glu Lys Gly Phe Val Gly Pro Tyr His Ser Lys Glu Met Pro His Leu
166      485      490      495
169 Val Asn Leu Ser Asp Asp Pro Leu Leu Ala Glu Cys Leu Val Tyr Asn

```

RAW SEQUENCE LISTING

DATE: 09/16/2003

PATENT APPLICATION: US/09/724,586A

TIME: 16:00:02

Input Set : A:\04871~1.txt

Output Set: N:\CRF4\09162003\I724586A.raw

```

170          500          505          510
173 Ile Lys Pro Gly Gln Thr Arg Val Gly Asn Val Asn Gln Asp Thr Gln
174          515          520          525
177 Ala Glu Ile Arg Leu Asn Gly Ser Lys Ile Leu Lys Glu His Cys Thr
178          530          535          540
181 Phe Glu Asn Val Asp Asn Val Val Thr Ile Val Pro Asn Glu Lys Ala
182 545          550          555          560
185 Ala Val Met Val Asn Gly Val Arg Ile Asp Lys Pro Thr Arg Leu Arg
186          565          570          575
189 Ser Gly Tyr Arg Ile Ile Leu Gly Asp Phe His Ile Phe Arg Phe Asn
190          580          585          590
193 His Pro Glu Glu Ala Arg Ala Glu Arg Gln Glu Gln Ser Leu Leu Arg
194          595          600          605
197 His Ser Val Thr Asn Ser Gln Leu Gly Ser Pro Ala Pro Gly Arg His
198          610          615          620
201 Asp Arg Thr Leu Ser Lys Ala Gly Ser Asp Ala Asp Gly Asp Ser Arg
202 625          630          635          640
205 Ser Asp Ser Pro Leu Pro His Phe Arg Gly Lys Asp Ser Asp Trp Phe
206          645          650          655
209 Tyr Ala Arg Arg Glu Ala Ala Ser Ala Ile Leu Gly Leu Asp Gln Lys
210          660          665          670
213 Ile Ser His Leu Thr Asp Asp Glu Leu Asp Ala Leu Phe Asp Asp Val
214          675          680          685
217 Gln Lys Ala Arg Ala Val Arg Arg Gly Leu Val Glu Asp Asn Glu Asp
218          690          695          700
221 Ser Asp Ser Gln Ser Ser Phe Pro Val Arg Asp Lys Tyr Met Ser Asn
222 705          710          715          720
225 Gly Thr Ile Asp Asn Phe Ser Leu Asp Thr Ala Ile Thr Met Pro Gly
226          725          730          735
229 Thr Pro Arg Ser Asp Asp Asp Gly Asp Ala Leu Phe Phe Gly Asp Lys
230          740          745          750
233 Lys Ser Lys Gln Asp Ala Ser Asn Val Asp Val Glu Glu Leu Arg Gln
234          755          760          765
237 Gln Gln Ala Gln Met Glu Glu Ala Leu Lys Thr Ala Lys Gln Glu Phe
238          770          775          780

```

241 <210> SEQ ID NO: 2

242 <211> LENGTH: 2352

243 <212> TYPE: DNA

244 <213> ORGANISM: Thermomyces lanuginosus

247 <220> FEATURE:

248 <221> NAME/KEY: misc_feature

249 <223> OTHER INFORMATION: TL-gamma ATP-dependent plus end-directed microtubule motor protein

252 <400> SEQUENCE: 2

```

253 atgtcggggtg gtggaatat caaggtggtg gtgcgggtac gcccgttcaa cgcccgagaa      60
255 atcgaccgtg gcgcaaatg tattgtgcgg atggaaggaa atcaaaccat cctcaccct      120
257 cctccgggtg ccgaagagaa ggcgcgtaaa agtggcaaaa ctattatgga tggcccgag      180
259 gcatttgcgt tcgatcggtc gtattggtcc ttgacaaga atgctcccaa ctatgcgaga      240
261 caggaagacc tattccaaga tctcggagtc ccgcttctgg ataatgcatt caagggttat      300

```

RAW SEQUENCE LISTING

DATE: 09/16/2003

PATENT APPLICATION: US/09/724,586A

TIME: 16:00:02

Input Set : A:\04871~1.txt

Output Set: N:\CRF4\09162003\I724586A.raw

```

263 aacaattgta tcttcgccta cggtcagacc ggttcgggca agtcctattc aatgatgggc 360
265 tatggcaagg agcatggcgt gatcccgcg atttgccagg acatgttccg gcgtattaat 420
267 gaactgcaga aggacaagaa cctcacttgc accgtcgaag tttcgtactt ggaaatttac 480
269 aatgaacgag tgcgagactt gctgaatccg tcgacaaagg ggaatctcaa ggtccgagaa 540
271 caccgcgcga ccggccccta cgtggaggac ttggcgaagc tggtcgtgcg atcattccaa 600
273 gaaatcgaaa atctcatgga tgagggcaac aaagccagaa cggttgccgc cacaacatg 660
275 aacgagacat ccagtcgatc ccacgccgtc ttcactttga ccttgacgca aaagtggcat 720
277 gatgaagaga ccaaaatgga cacagagaag gttgcgaaga tcagtctggt agatttggcg 780
279 ggttctgagc gagcaacgtc caccggagct actggagcgc gactgaagga ggggtgcagag 840
281 atcaaccgct cactttcgac cctaggtcgt gtgattgcag cgctagcgga tatgtcgtcg 900
283 ggaaaacaga agaagaatca gttagtacct taccgagatt cggtagtgac gtggcttctg 960
285 aaggactcct tgggaggcaa ctcgatgacc gccatgattg ccgccatttc gcctgctgat 1020
287 attactttg aagagactct cagtaccctt cgatatgcgg actctgcgaa gcgaatcaag 1080
289 aaccacgcag tggtaaatga agaccggaac gcgcggatga tccgcgagtt gaaggaggaa 1140
291 ctgcgcgagc tgaggagcaa actccagagc agtggtaggag gtggagggtg tgcaggaggt 1200
293 tctggcgggc cagtggagga atcgtaccgc ccgcacacgc cgctcgagaa gcaaatcgtg 1260
295 tcgattcagc agccggatgc gacagtcaag aaaatgagca aggcagaaat cgtggagcaa 1320
297 ctgaaccaga gtgagaagct ctatcgggat ctcaatcaga cctgggaaga gaagctggcc 1380
299 aagaccgagg aaatccacaa ggaacgagaa gcggcgctcg aggagctggg tatcagcatc 1440
301 gaaaagggct ttgttgccc ttaccactcc aaagaaatgc cacatctagt caacttgagc 1500
303 gatgacctc ttctggctga gtgtcttgc tacaacatca agcccgggca gacaagggtt 1560
305 ggaacgtca accaagatac acaagcgga attcgtctga acggttcgaa gatcctgaaa 1620
307 gaacactgta cgtttgaaaa tgtggacaac gttgtgacca tcgtgcaaaa cgagaaggct 1680
309 gctgtcatgg tgaacggcgt gcgaatcgac aagcctact gcctccgag cggtacagg 1740
311 atcatcctgg gcgatttcca catttttcga ttcaaccatc cggaagaagc tcgtgcggaa 1800
313 cggcaagaac aatccttgct tcgccattct gtcaccaaca gtcagttggg ttcgcctgct 1860
315 ccaggccgtc acgaccggac actgagcaag gcgggttcgg atgcggacgg cgattctcgc 1920
317 tcagattctc ctttgccgca ctttcgtgga aaggatagcg actggttcta tgctcgcagg 1980
319 gaagctgcta gcgcgatcct agggttggat cagaagatct ctcatctgac agatgacgag 2040
321 ttggatgcat tatttgacga tgttcagaaa gcgcgggcag ttcgtcgtgg gctggtcgaa 2100
323 gacaacgaag atagcgattc gcagagttcg tttccggtcc gtgacaaata catgtccaat 2160
325 ggaaccattg ataatttctc gctcgatacc gccattacta tgccgggtac ccctcgtagt 2220
327 gatgacgacg gtgacgcgct gttttttggt gataagaagt cgaaacagga tgcgtctaata 2280
329 gttgatgttg aggagttgag tcaacagcag gctcagatgg aagaagccct gaaaacagcg 2340
331 aagcaggaat tc 2352
334 <210> SEQ ID NO: 3
335 <211> LENGTH: 21
336 <212> TYPE: DNA
337 <213> ORGANISM: Artificial Sequence
339 <220> FEATURE:
340 <223> OTHER INFORMATION: Synthetic
342 <400> SEQUENCE: 3
343 atgtcgggcg gtggaaatat c 21
346 <210> SEQ ID NO: 4
347 <211> LENGTH: 23
348 <212> TYPE: DNA
349 <213> ORGANISM: Artificial Sequence
351 <220> FEATURE:
352 <223> OTHER INFORMATION: Synthetic

```

RAW SEQUENCE LISTING

DATE: 09/16/2003

PATENT APPLICATION: US/09/724,586A

TIME: 16:00:02

Input Set : A:\04871~1.txt

Output Set: N:\CRF4\09162003\I724586A.raw

```

354 <400> SEQUENCE: 4
355 gaattcctgc ttcgctgttt tca 23
358 <210> SEQ ID NO: 5
359 <211> LENGTH: 30
360 <212> TYPE: DNA
361 <213> ORGANISM: Artificial Sequence
363 <220> FEATURE:
364 <223> OTHER INFORMATION: Synthetic
367 <220> FEATURE:
368 <221> NAME/KEY: misc_feature
369 <222> LOCATION: (25)..(25)
370 <223> OTHER INFORMATION: n is a, c, g, or t
372 <400> SEQUENCE: 5
W--> 373 gcgcgggatcc atytttygcht ayggncarac 30
376 <210> SEQ ID NO: 6
377 <211> LENGTH: 30
378 <212> TYPE: DNA
379 <213> ORGANISM: Artificial Sequence
381 <220> FEATURE:
382 <223> OTHER INFORMATION: Synthetic
385 <220> FEATURE:
386 <221> NAME/KEY: misc_feature
387 <222> LOCATION: (16)..(16)
388 <223> OTHER INFORMATION: n is a, c, g, or t
390 <220> FEATURE:
391 <221> NAME/KEY: misc_feature
392 <222> LOCATION: (28)..(28)
393 <223> OTHER INFORMATION: n is a, c, g, or t
395 <400> SEQUENCE: 6
W--> 396 gcgcgaattc tcdganccdg cvarrtcnac 30
399 <210> SEQ ID NO: 7
400 <211> LENGTH: 30
401 <212> TYPE: DNA
402 <213> ORGANISM: Artificial Sequence
404 <220> FEATURE:
405 <223> OTHER INFORMATION: Synthetic
408 <220> FEATURE:
409 <221> NAME/KEY: misc_feature
410 <222> LOCATION: (16)..(16)
411 <223> OTHER INFORMATION: n is a, c, g, or t
413 <220> FEATURE:
414 <221> NAME/KEY: misc_feature
415 <222> LOCATION: (28)..(28)
416 <223> OTHER INFORMATION: n is a, c, g, or t
418 <400> SEQUENCE: 7
W--> 419 gcgcgaattc tcdctnccdg cvarrtcnac 30
422 <210> SEQ ID NO: 8
423 <211> LENGTH: 21
424 <212> TYPE: DNA

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/724,586A

DATE: 09/16/2003
TIME: 16:00:03

Input Set : A:\04871~1.txt
Output Set: N:\CRF4\09162003\I724586A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; N Pos. 25
Seq#:6; N Pos. 16,28
Seq#:7; N Pos. 16,28